

ECON 523: Program Evaluation for International Development

Empirical Exercise 5

In this exercise, we'll be using a data set on primary school enrollment and completion in 35 African countries, 15 of which eliminated primary school fees between 1990 and 2015. Raw data on country-level enrollment and completion rates comes from the World Bank's World Development Indicators Database. We will use this data set to generate two-way fixed effects (TWFE) estimates of the impact of eliminating school fees on primary school completion. Since this policy was phased in by different countries at different times, it is a useful setting for exploring the strengths and weaknesses of two-way fixed effects.

Before you begin, create a do file (with all the standard stuff at the top) that downloads the data set `E5-fpdata.dta` from the course website. Drop any country-years that are missing data on the primary school completion rate. Extend your do file as you answer the following questions:

1. Estimate two TWFE regressions of primary school completion on `fpe` controlling for country and year fixed effects. In your first regression, include only the countries that eventually implemented free primary; include all the countries in the data set in your second regression. Cluster your standard errors at the country level. Export your regression results to word, and take a screen shot of your resulting (nice looking) table.
2. Given what you know about TWFE, and given the nature of the policy and outcome under consideration, why might you have expected the coefficient (on `fpe`) in Column 1 to be smaller than the coefficient in Column 2?
3. **Negative weights.**
 - (a) What proportion of treated country years receive negative weighting in the TWFE estimation when you exclude the never treated countries? You can answer this question by generating and then summarizing a variable that is 1 for treated observations (i.e. country-years) receiving negative weight in TWFE, 0 for treated observations receiving positive weight in TWFE, and missing for untreated country-years.
 - (b) What proportion of treated country years receive negative weighting in the TWFE estimation when you include the never treated countries?
4. **Event studies.**
 - (a) Generate a relative time variable `rel_time` that indicates the difference between the (current, for each observation) year and the year in which FPE was implemented in that country.
 - (b) Summarize the `rel_time` variable. What is the maximum number of years that we observe **before** a country implements free primary (among countries that eventually implement it)? Use the loop below to generate variables `minus_2`, `minus_3`, `minus_4` etc. that are dummies equal to one for country-years (respectively) 2, 3, 4 etc. years before a country implements free primary. For countries that never implement free primary, these variables should be equal to 0 for all years.

```

forvalues i = 2/27 {
gen minus_`i' = (fpe_year - year==`i' & nevertreated==0)
}

```

- (c) What is the maximum number of years that we observe **after** a country implements free primary (among countries that eventually implement it)? Generate variables `plus_0`, `plus_1`, `plus_2` etc that are dummies equal to one for country-years (respectively) 0, 1, 2 etc. years after a country implements free primary. The variable `plus_0` indicates the year FPE was first implemented. For countries that never implement free primary, these variables should be equal to 0 for all years.
- (d) Now implement the event study design by regressing complete on country and year fixed effects as well as the `minus_*` and `plus_*` variables. What patterns of significance do you observe among the `plus_*` variables? Is there ever a statistically significant impact of FPE on primary school completion? What patterns of significance do you observe among the `minus_*` variables? Is there evidence that the assumption of common trends is violated?
- (e) Adapt the code from the do file `ECON523-E5-event-study-example.do` to make an event study graph of your results. Save the graph as a `pdf` or `png` file (so that you can upload it later). What does the graph suggest about your TWFE model?

5. Trimming.

- (a) Rerun your event study regressions. Use the post-estimation `test` command to test whether you can reject the hypothesis that `minus_2` is equal to 0. (You should not be able to). Now test the joint hypothesis that `minus_2` and `minus_3` are both equal to 0. Continue adding additional pre-treatment terms to your joint test until you can reject the hypothesis that they are all zero at the 90 percent level.
 - (b) Use the results from Question 5a to exclude some pre-treatment observations (where common trends appears to be violated). In other words, drop observations based on their relative time, so that you can implement an event study design where none of your pre-FPE terms is statistically significant.
 - (c) Make a new event study plot that presents the results in your trimmed (i.e. restricted sample).
 - (d) Make a table that reports the results of two TWFE regressions of primary school completion rates on the FPE dummy. In Column 1, include only the countries that eventually implemented free primary. In Column 2, include all the countries in the sample. Cluster your standard errors at the country level, report standard errors rather than t-statistics, and make your table look as professional as possible. Export your results to word, and take a screen shot of your finished table (so that you can upload it later).
6. From a statistical perspective, do you think it is good or bad that we trimmed the data by including only eight years prior to the implementation of free primary? Justify your answer (in no more than 2 or 3 sentences).